

Contents lists available at [SciVerse ScienceDirect](http://SciVerse.ScienceDirect.com)

## Theoretical Computer Science

journal homepage: [www.elsevier.com/locate/tcs](http://www.elsevier.com/locate/tcs)

# New techniques for approximating optimal substructure problems in power-law graphs

Yilin Shen<sup>\*</sup>, Dung T. Nguyen, Ying Xuan, My T. Thai

Department of Computer Information Science and Engineering, University of Florida, Gainesville, FL, 32611, United States

## ARTICLE INFO

### Keywords:

Power-law graphs  
Optimal substructure problems  
Embedding techniques  
Inapproximability optimal substructure framework  
Complexity hardness  
Inapproximability

## ABSTRACT

The remarkable discovery of many large-scale real networks is the power-law distribution in degree sequence: the number of vertices with degree  $i$  is proportional to  $i^{-\beta}$  for some constant  $\beta > 1$ . A lot of researchers believe that it may be easier to solve some optimization problems in power-law graphs. Unfortunately, many problems have been proved *NP*-hard even in power-law graphs. Intuitively, a theoretical question is raised: are these problems on power-law graphs still as hard as on general graphs?

In this paper, we show that many optimal substructure problems, such as MINIMUM DOMINATING SET, MINIMUM VERTEX COVER and MAXIMUM INDEPENDENT SET, are easier to solve in power-law graphs by illustrating better inapproximability factors. An optimization problem has the property of optimal substructure if its optimal solution on some given graph is essentially the union of the optimal sub-solutions on all maximal connected components. In particular, we prove the above problems and a more general problem ( $\rho$ -MINIMUM DOMINATING SET) remain *APX*-hard and their constant inapproximability factors on general power-law graphs by using the cycle-based embedding technique to embed any  $d$ -bounded graphs into a power-law graph. In addition, in simple power-law graphs, we further prove the corresponding inapproximability factors of these problems based on the graphic embedding technique as well as that of MAXIMUM CLIQUE and MINIMUM COLORING using the embedding technique in [1]. As a result of these inapproximability factors, the belief that there exists some  $(1 + o(1))$ -approximation algorithm for these problems on power-law graphs is proven to be not always true. At last, we do in-depth investigations in the relationship between the exponential factor  $\beta$  and constant greedy approximation algorithms.

© 2011 Elsevier B.V. All rights reserved.

## 1. Introduction and related work

A great number of large-scale networks in real life are discovered to follow a power-law distribution in their degree sequences, ranging from the Internet [2], the World-Wide Web (WWW) [3] to social networks [4]. That is, the number of vertices with degree  $i$  is proportional to  $i^{-\beta}$  for some constant  $\beta$  in these graphs, which is called power-law graphs. The observations show that the exponential factor  $\beta$  ranges between 1 and 4 for most real-world networks [5]. Intuitively, the following theoretical question is raised: what are the differences in terms of complexity hardness and inapproximability factor of several optimization problems between in general graphs and in power-law graphs?

Many experimental results on random power-law graphs give us a belief that the problems might be much easier to solve on power-law graphs. Eubank et al. [6] showed that a simple greedy algorithm leads to a  $1 + o(1)$  approximation factor on MINIMUM DOMINATING SET (MDS) and MINIMUM VERTEX COVER (MVC) on power-law graphs (without any

<sup>\*</sup> Corresponding author. Tel.: +1 352 226 9666.

E-mail addresses: [yshen@cise.ufl.edu](mailto:yshen@cise.ufl.edu) (Y. Shen), [dtnguyen@cise.ufl.edu](mailto:dtnguyen@cise.ufl.edu) (D.T. Nguyen), [yxuan@cise.ufl.edu](mailto:yxuan@cise.ufl.edu) (Y. Xuan), [mythai@cise.ufl.edu](mailto:mythai@cise.ufl.edu) (M.T. Thai).

**Table 1**Inapproximability factors on power-law graphs with exponential factor  $\beta > 1$ .

Problem	General power-law graph	Simple power-law graph
MIS	$1 + \frac{1}{140(2\zeta(\beta)3^\beta - 1)} - \varepsilon$	$1 + \frac{1}{1120\zeta(\beta)3^\beta} - \varepsilon$
MDS	$1 + \frac{1}{390(2\zeta(\beta)3^\beta - 1)}$	$1 + \frac{1}{3120\zeta(\beta)3^\beta}$
MVC, $\rho$ -MDS	$1 + \frac{2(1 - (2 + o_c(1)) \frac{\log \log c}{\log c})}{\left(\zeta(\beta)c^\beta + c^{\frac{1}{\beta}}\right)(c+1)}$	$1 + \frac{2 - (2 + o_c(1)) \frac{\log \log c}{\log c}}{2\zeta(\beta)c^\beta(c+1)}$
CLIQUE	–	$O\left(n^{1/(\beta+1)-\epsilon}\right)$
COLORING	–	$O\left(n^{1/(\beta+1)-\epsilon}\right)$

<sup>a</sup> Conditions: MIS and MDS:  $P \neq NP$ ; MVC,  $\rho$ -MDS: unique games conjecture; CLIQUE, COLORING:  $NP \neq ZPP$ .

<sup>b</sup>  $c$  is a constant which is the smallest  $d$  satisfying the condition in [15].

formal proof) although MDS and MVC has been proved  $NP$ -hard to be approximated within  $(1 - \epsilon) \log n$  and 1.366 on general graphs respectively [7]. In [8], Gopal also claimed that there exists a polynomial time algorithm that guarantees a  $1 + o(1)$  approximation of the MVC problem with probability at least  $1 - o(1)$ . Unfortunately, there is no such formal proof for this claim either. Furthermore, several papers also have some theoretical guarantees for some problems on power-law graphs. Gkantsidis et al. [9] proved the flow through each link is at most  $O(n \log^2 n)$  on power-law random graphs where the routing of  $O(d_u d_v)$  units of flow between each pair of vertices  $u$  and  $v$  with degrees  $d_u$  and  $d_v$ . In [9], the authors take advantage of the property of power-law distribution by using the structural random model [10,11] and show the theoretical upper bound with high probability  $1 - o(1)$  and the corresponding experimental results. Likewise, Janson et al. [12] gave an algorithm that approximated MAXIMUM CLIQUE within  $1 - o(1)$  on power-law graphs with high probability on the random Poisson model  $G(n, \alpha)$  (i.e. the number of vertices with degree at least  $i$  decreases roughly as  $n^{-i}$ ). Although these results were based on experiments and various random models, they raise an interest in investigating hardness and inapproximability of optimization problems on power-law graphs.

Recently, Ferrante et al. [1] had an initial attempt on power-law graphs to show the  $NP$ -hardness of MAXIMUM CLIQUE (CLIQUE) and MINIMUM GRAPH COLORING (COLORING) ( $\beta > 1$ ) by constructing a bipartite graph to embed a general graph into a power-law graph and  $NP$ -hardness of MVC, MDS and MAXIMUM INDEPENDENT SET (MIS) ( $\beta > 0$ ) based on their optimal substructure properties. Unfortunately, there is a minor flaw in the proof of their Lemma 5 which makes the proof of  $NP$ -hardness of MIS, MVC, MDS with  $\beta < 1$  no longer hold. Then we present another way in the Appendix to show the  $NP$ -hardness of these problems when  $\beta < 1$  so as to fix this non-trivial flaw.

*Our contributions.* In this paper, we propose two new techniques on optimal substructure problems, *Cycle-Based Embedding Technique* and *Graphic Embedding Technique*, to embed a  $d$ -bounded graph into a general power-law graph and a simple power-law graph respectively. Then we use these two techniques to further prove the  $APX$ -hardness and the inapproximability of MIS, MDS, and MVC on general power-law graphs and simple power-law graphs. These inapproximability results on power-law graphs are shown in Table 1. Furthermore, the inapproximability results in CLIQUE and COLORING are shown by taking advantage of the reduction in [1]. We also analyze the relationship between  $\beta$  and the constant greedy approximation algorithms for MIS and MDS.

In addition, due to a lot of recent studies in online social networks on the influence propagation problem [13,14], we formulate this problem as  $\rho$ -Minimum Dominating Set ( $\rho$ -MDS) and show it hard to be approximated within  $2 - (2 + o_d(1)) \log \log d / \log d$  factor on  $d$ -bounded graphs under unique games conjecture, which further leads to the following inapproximability result on power-law graphs (shown in Table 1).

The rest of paper is organized as follows. In Section 2, we introduce some problem definitions, the model of power-law graphs, and some related concepts. The inapproximability optimal substructure framework is presented in Section 3. We show the hardness and inapproximability of MIS, MDS, MVC in general power-law graphs using the cycle-based embedding technique in Section 4. More inapproximability results in simple power-law graphs are illustrated in Section 5 based on the graphic embedding technique, which implies the  $APX$ -hardness of these problems. Additionally, the inapproximability factor on maximum clique and minimum coloring problems are proven. In Section 6, we analyze the relationship between  $\beta$  and constant approximation algorithms, which further proves that the integral gap is typically small for optimization problems on power-law graphs than that on general bounded graphs. In the Appendix, we fix the flaw in the  $NP$ -hardness proof for  $\beta < 1$  presented in [1].

## 2. Preliminaries

In this section, we first recall the definition of several classical optimization problems and formulate the new optimization problem  $\rho$ -Minimum Dominating Set. Then the power-law model and some corresponding concepts are proposed. At last, we introduce some special graphs which will be used in the analysis throughout the whole paper.

## 2.1. Problem definitions

**Definition 2.1** (*Maximum Independent Set*). Given an undirected graph  $G = (V, E)$ , find a subset  $S \subseteq V$  with the maximum size such that no two vertices in  $S$  are adjacent.

**Definition 2.2** (*Minimum Vertex Cover*). Given an undirected graph  $G = (V, E)$ , find a subset  $S \subseteq V$  with the minimum size such that for each edge  $E$  at least one endpoint belongs to  $S$ .

**Definition 2.3** (*Minimum Dominating Set*). Given an undirected graph  $G = (V, E)$ , find a subset  $S \subseteq V$  with the minimum size such that for each vertex  $v_i \in V \setminus S$ , at least one neighbor of  $v_i$  belongs to  $S$ .

**Definition 2.4** (*Maximum Clique*). Given an undirected graph  $G = (V, E)$ , find a clique with maximum size where a subgraph of  $G$  is called a clique if all its vertices are pairwise adjacent.

**Definition 2.5** (*Minimum Graph Coloring*). Given an undirected graph  $G = (V, E)$ , label the vertices in  $V$  with minimum number of colors such that no two adjacent vertices share the same color.

The  $\rho$ -Minimum Dominating Set is a general version of MDS problem. In the context of influence propagation, the  $\rho$ -MDS problem aims to find a subset of nodes with minimum size such that all nodes in the whole network can be influenced within  $t$  rounds. In particular, a node is influenced when  $\rho$  fraction of its neighbors are influenced. For simplicity, we define  $\rho$ -MDS problem in the case that  $t = 1$ .

**Definition 2.6** ( $\rho$ -Minimum Dominating Set). Given an undirected graph  $G = (V, E)$ , find a subset  $S \subseteq V$  with the minimum size such that for each vertex  $v_i \in V \setminus S$ ,  $|S \cap N(v_i)| \geq \rho |N(v_i)|$ .

## 2.2. Power-law model and some notations

A great number of models [16,17,10,11,18] on power-law graphs are emerging in the past recent years. In this paper, we do the analysis based on the general  $(\alpha, \beta)$  model, that is, the graphs only constrained by the power-law distribution in degree sequences. We first define the following two types of degree sequences.

**Definition 2.7** (*y-Degree Sequence*). Given a graph  $G = (V, E)$ , the  $y$ -degree sequence of  $G$  is a sequence  $Y = \langle y_1, y_2, \dots, y_\Delta \rangle$  where  $\Delta$  is the maximum degree of  $G$  and  $y_i = |\{u | u \in V \wedge \deg(u) = i\}|$ .

**Definition 2.8** (*d-Degree Sequence*). Given a graph  $G = (V, E)$ , the  $d$ -degree sequence of  $G$  is a sequence  $D = \langle d_1, d_2, \dots, d_n \rangle$  of vertex in non-increasing order of their degrees.

Note that  $y$ -degree sequence and  $d$ -degree sequence are interchangeable. Given a  $y$ -degree sequence  $Y = \langle y_1, y_2, \dots, y_\Delta \rangle$ , the corresponding  $d$ -degree sequence is  $D = \langle \Delta, \Delta, \dots, \Delta - 1, \Delta - 1, \dots, \Delta - 1, \dots, 1, \dots, 1 \rangle$  where the number  $i$  appears  $y_i$  times. Because of their equivalence, we may use only  $y$ -degree sequence or  $d$ -degree sequence or both without changing the meaning or validity of results. The definition of power-law graphs can be expressed via  $y$ -degree sequences as follows.

**Definition 2.9** (*General  $(\alpha, \beta)$  Power-Law Graph Model*). A graph  $G = (V, E)$  is called a  $(\alpha, \beta)$  power-law graph  $G_{(\alpha, \beta)}$  where multi-edges and self-loops are allowed if the maximum degree is  $\Delta = \lfloor e^{\alpha/\beta} \rfloor$  and the number of vertices of degree  $i$  is:

$$y_i = \begin{cases} \lfloor e^\alpha / i^\beta \rfloor, & \text{if } i > 1 \text{ or } \sum_{i=1}^{\Delta} \lfloor e^\alpha / i^\beta \rfloor \text{ is even} \\ \lfloor e^\alpha \rfloor + 1, & \text{otherwise.} \end{cases} \quad (2.1)$$

In simple  $(\alpha, \beta)$  power-law graphs, there are no multi-edges and self-loops.

Note that a power-law graph are represented by two parameters  $\alpha$  and  $\beta$ . Since graphs with the same  $\beta$  exhibit the same behaviors, we categorize all graphs with the same  $\beta$  into a  $\beta$ -family of graphs such that  $\beta$  is regarded as a constant instead of an input. In addition, we only consider the case  $\beta > 1$  because almost all real large-scale networks have  $\beta > 1$ . In this case, the number of vertices is:

$$\sum_{i=1}^{\Delta} \frac{e^\alpha}{i^\beta} = \zeta(\beta) e^\alpha - O\left(n^{\frac{1}{\beta}-1}\right) \approx \zeta(\beta) e^\alpha$$

where  $\zeta(\beta) = \sum_{i=1}^{\infty} \frac{1}{i^\beta}$  is the Riemann Zeta Function. Also the  $d$ -degree sequence of any  $(\alpha, \beta)$  power-law graph is continuous according to the following definition.

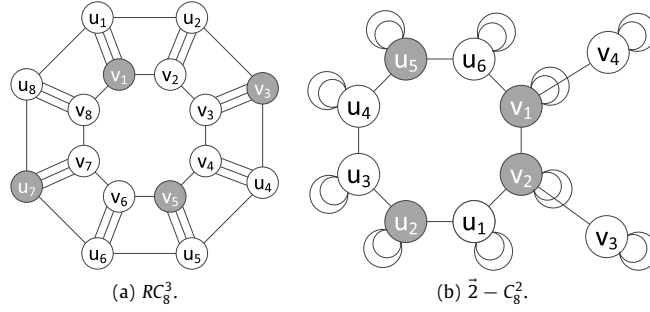
**Definition 2.10** (*Continuous Sequence*). An integer sequence  $\langle d_1, d_2, \dots, d_n \rangle$ , where  $d_1 \geq d_2 \geq \dots \geq d_n$ , is continuous if  $\forall 1 \leq i \leq n-1, |d_i - d_{i+1}| \leq 1$ .

**Definition 2.11** (*Graphic Sequence*). A sequence  $D$  is said to be graphic if there exists a graph such that  $D$  is its  $d$ -degree sequence.

**Definition 2.12** (*Degree Set*). Given a graph  $G$ , let  $D_i(G)$  be the set of vertices of degree  $i$  on  $G$ .

Furthermore, we define the  $d$ -bounded graph as

**Definition 2.13** (*d-Bounded Graph*). Given a graph  $G = (V, E)$ ,  $G$  is a  $d$ -bounded graph if the degree of any vertex is upper bounded by an integer constant  $d$ .



**Fig. 1.** Special graph examples: the left one is a  $(3, 3, 3, 3, 3, 3, 3, 3)$ -regular cycle and the right one is a  $(3, 3, 3, 3)$ -branch- $(2, 2, 2, 2, 2, 2)$ -cycle. The gray vertices consist of the optimal solution of MDS on these two special graphs.

### 2.3. Special graphs

**Definition 2.14** ( $\vec{d}$ -Regular Cycle  $RC_n^{\vec{d}}$ ). Given a vector  $\vec{d} = (d_1, \dots, d_n)$ , a  $\vec{d}$ -regular cycle  $RC_n^{\vec{d}}$  is composed of two cycles. Each cycle has  $n$  vertices and two  $i^{\text{th}}$  vertices in each cycle are adjacent with each other by  $d_i - 2$  multi-edges. That is,  $\vec{d}$ -regular cycle  $RC_n^{\vec{d}}$  has  $2n$  vertices and the two  $i^{\text{th}}$  vertices have the same degree  $d_i$ . An example  $RC_8^3$  is shown in Fig. 1(a).

**Definition 2.15** ( $\vec{\kappa}$ -Branch- $\vec{d}$ -Cycle  $\vec{\kappa}\text{-}BC_n^{\vec{d}}$ ). Given two vectors  $\vec{d} = (d_1, \dots, d_n)$  and  $\vec{\kappa} = (\kappa_1, \dots, \kappa_m)$ , the  $\vec{\kappa}$ -branch- $\vec{d}$ -cycle is composed of a cycle with a number of vertices  $n$  such that each vertex has degree  $d_i$  as well as  $|\vec{\kappa}|/2$  appendant branches, where  $|\vec{\kappa}|$  is an even number. Note that any  $\vec{\kappa}$ -branch- $\vec{d}$ -cycle has  $|\vec{\kappa}|$  even number of vertices with odd degrees. An example is shown in Fig. 1(b).

### 2.4. Existing inapproximability results

Here we list some inapproximability results in the literature to use later in our proofs.

- (1) In  $d$ -bounded graphs, MVC is hard to be approximated into  $2 - (2 + o_d(1)) \log \log d / \log d$  for every sufficiently large integer  $d$  under unique games conjecture [15,19].
- (2) In 3-bounded graphs, MIS and MDS is NP-hard to be approximated into  $\frac{140}{139} - \varepsilon$  for any  $\varepsilon > 0$  and  $\frac{391}{390}$  respectively [20].
- (3) Maximum clique and minimum coloring problem is hard to be approximated into  $n^{1-\epsilon}$  on general graphs unless NP=ZPP [21].

## 3. Inapproximability optimal substructure framework in power-law graphs

In this section, we introduce a framework to derive the approximation hardness of optimal substructure problems on power-law graphs. A graph optimization problem is said to satisfy optimal substructure if its optimal solution is the union of the optimal solutions on each connected component. Therefore, when a graph  $G$  is embedded into a power-law graph  $G'$ , the optimal solution in  $G'$  consists of a subset of the optimal solution in  $G$ . According to this important property, we present the *Inapproximability Optimal Substructure Framework* to prove the inapproximability factor if there exists an *Embedded-Approximation-Preserving Reduction* that relates the approximation hardness in general graphs and power-law graphs by guaranteeing the relationship between the solutions in the original graph and the constructed graph.

**Definition 3.1** (*Embedded-Approximation-Preserving Reduction*). Given an optimal substructure problem  $O$ , a reduction from an instance on graph  $G = (V, E)$  to another instance on a power-law graph  $G' = (V', E')$  is called embedded-approximation-preserving if it satisfies the following properties:

- (1)  $G$  is a subset of maximal connected components of  $G'$ ;
- (2) The optimal solution of  $O$  on  $G'$ ,  $OPT(G')$ , is upper bounded by  $\mathfrak{C}OPT(G)$  where  $\mathfrak{C}$  is a constant correspondent to the growth of the optimal solution.

**Theorem 3.1** (*Inapproximability Optimal Substructure Framework*). Given an optimal substructure problem  $O$ , if there exists an embedded-approximation-preserving reduction from a graph  $G$  to another graph  $G'$ , we can extract the inapproximability factor  $\delta$  of  $O$  on  $G'$  using  $\epsilon$ -inapproximability of  $O$  on  $G$ , where  $\delta$  is lower bounded by  $\frac{\epsilon \mathfrak{C}}{(\mathfrak{C}-1)\epsilon+1}$  and  $\frac{\epsilon+\mathfrak{C}-1}{\mathfrak{C}}$  when  $O$  is a maximum and minimum optimization problem respectively.

**Proof.** Suppose that there exists an algorithm providing a solution of  $O$  on  $G'$  with size at most  $\delta$  times the optimal solution. Denote  $A$  and  $B$  to be the sizes of the produced solution on  $G$  and  $G' \setminus G$  and  $A^*$  and  $B^*$  to be their corresponding optimal

values. Hence, we have  $B^* \leq (\mathfrak{C} - 1)A^*$ . With the completeness that  $OPT(G) = A^* \Rightarrow OPT(G') = B^*$ , the soundness leads to the lower bound of  $\delta$  which is dependent on the type of  $O$ , maximization or minimization problem, as follows.

**Case 1:** When  $O$  is a maximization problem, we start from the definition of soundness as

$$A^* + B^* \leq \delta(A + B) \quad (3.1)$$

$$\Leftrightarrow A^* \leq \delta A + (\delta - 1)B^* \quad (3.2)$$

$$\Leftrightarrow A^* \leq \delta A + (\delta - 1)(\mathfrak{C} - 1)A^* \quad (3.3)$$

where (3.2) holds since  $B \leq B^*$  and (3.3) holds since  $B^* \leq (\mathfrak{C} - 1)A^*$ .

On the other hand, it is hard to approximate  $O$  within  $\epsilon$  on  $G$ , thus  $A^* > \epsilon A$ . Replace it to the above inequality, we have:

$$A^* < A^* \delta / \epsilon + (\delta - 1)(\mathfrak{C} - 1)A^* \Leftrightarrow \delta > \frac{\epsilon \mathfrak{C}}{(\mathfrak{C} - 1)\epsilon + 1}$$

**Case 2:** When  $O$  is a minimization problem, since  $B^* \leq B$ , similarly

$$A + B \leq \delta(A^* + B^*)$$

$$\Leftrightarrow A \leq \delta A^* + (\delta - 1)B^*$$

$$\Leftrightarrow A \leq \delta A^* + (\delta - 1)(\mathfrak{C} - 1)A^*$$

Then from  $A > \epsilon A^*$ ,

$$\epsilon < \delta + (\delta - 1)(\mathfrak{C} - 1) \Leftrightarrow \delta > \frac{\epsilon + \mathfrak{C} - 1}{\mathfrak{C}}. \quad \square$$

## 4. Hardness and inapproximability of optimal substructure problems on general power-law graphs

### 4.1. General Cycle-Based Embedding Technique

In this section, we propose a *General Cycle-Based Embedding Technique* on  $(\alpha, \beta)$  power-law graphs with  $\beta > 1$ . The basic idea is to embed an arbitrary  $d$ -bounded graph into power-law graphs using a  $\vec{d}_1$ -regular cycle, a  $\vec{\kappa}$ -branch- $\vec{d}_2$ -cycle and a number of cliques  $K_2$ , where  $\vec{d}_1$ ,  $\vec{d}_2$  and  $\vec{\kappa}$  are defined by  $\alpha$  and  $\beta$ . Before discussing the main embedding technique, we first show that most optimal substructure problems can be polynomially solved in both  $\vec{d}$ -regular cycles and  $\vec{\kappa}$ -branch- $\vec{d}$ -cycle. In this context, the cycle-based embedding technique helps to prove the complexity of these optimal substructure problems on power-law graphs according to their corresponding complexity results on general bounded graphs.

**Lemma 4.1.** *MDS, MVC and MIS are polynomially solvable on  $\vec{d}$ -regular cycles.*

**Proof.** Here we just prove MDS problem can be polynomially solvable on  $\vec{d}$ -regular cycles. The algorithm is simple. From an arbitrarily vertex, we select the vertex on the other cycle in two hops. The algorithm will terminate until all vertices are dominated. Now we will show that this gives the optimal solution. Let us take  $RC_8^3$  as an example. As shown in Fig. 1(a), the size of MDS is 4. Notice that each vertex can dominate exact 3 vertices, that is, 4 vertices can dominate exactly 12 vertices. However, in  $RC_8^3$ , there are altogether 16 vertices, which have to be dominated by at least 4 vertices apart from the vertices in MDS. That is, the algorithm returns an optimal solution. The proof of MVC and MIS is similar.  $\square$

**Lemma 4.2.** *MDS, MVC and MIS can be polynomially solvable on  $\vec{\kappa}$ -branch- $\vec{d}$ -cycles.*

**Proof.** Again we show the proof of MDS. First we select the vertices connecting both the branches and the cycle. Then by removing the branches, we will have a line graph regardless of self-loops, on which MDS is polynomially solvable. It is easy to see that the size of MDS will increase if any one vertex connecting both the branch and the cycle in MDS is replaced by some other vertices. The proof of MIS is similar. Note that the optimal solution for MVC consists of all vertices since all edges need to be covered.  $\square$

**Theorem 4.1** (Cycle-Based Embedding Technique). *Any  $d$ -bounded graph  $G_d$  can be embedded into a power-law graph  $G_{(\alpha, \beta)}$  with  $\beta > 1$  such that  $G_d$  is a maximal component and most optimal substructure problems can be polynomially solvable on  $G_{(\alpha, \beta)} \setminus G_d$ .*

**Proof.** With the given  $\beta$ , we choose  $\alpha$  to be  $\max\{\ln \max_{1 \leq i \leq d}\{n_i \cdot i^\beta\}, \beta \ln d\}$ . Based on  $\tau(i) = \lfloor e^\alpha / i^\beta \rfloor - n_i$  where  $n_i = 0$  when  $i > d$ , we construct the power-law graph  $G_{(\alpha, \beta)}$  as the following Algorithm 1. The last step holds since the number of vertices of odd degrees has to be even. From Step 1, we know  $e^\alpha = \max\{\max_{1 \leq i \leq d}\{n_i \cdot i^\beta\}, d^\beta\} \leq d^\beta n$ , that is, the number of vertices  $N$  in graph  $G_{(\alpha, \beta)}$  satisfies  $N \leq \zeta(\beta)d^\beta n$ , which means that  $N/n$  is a constant. According to Lemmas 4.1 and 4.2, since  $G_{(\alpha, \beta)} \setminus G_d$  is composed of a  $\vec{d}_1$ -regular cycle and a  $\vec{d}_2$ -branch- $\vec{d}_2$ -cycle, it can be polynomially solvable. Note that the number of vertices in  $L$  is at most  $\Delta$  since there is at most one leftover vertex of each degree.  $\square$

**Algorithm 1:** Cycle Embedding Algorithm

- 
- 1  $\alpha \leftarrow \max\{\ln \max_{1 \leq i \leq d} \{n_i \cdot i^\beta\}, \beta \ln d\}$ ;
  - 2 For  $\tau(1)$  vertices of degree 1, add  $\lfloor \tau(1)/2 \rfloor$  number of cliques  $K_2$ ;
  - 3 For  $\tau(2)$  vertices of degree 2, add a cycle with the size  $\tau(2)$ ;
  - 4 For all vertices of degree larger than 2 and smaller than  $\Delta$ , construct a  $\vec{d}_1$ -regular cycle where  $\vec{d}_1$  is a vector composed of  $\lfloor \tau(i)/2 \rfloor$  number of elements  $i$  for all  $i$  satisfying  $\tau(i) > 0$ ;
  - 5 For all leftover isolated vertices  $L$  such that  $\tau(i) - 2\lfloor \tau(i)/2 \rfloor = 1$ , construct a  $\vec{d}_2^1$ -branch- $\vec{d}_2^2$ -cycle, where  $\vec{d}_2^1$  and  $\vec{d}_2^2$  are the vectors containing odd and even elements correspondent to the vertices of odd and even degrees in  $L$  respectively.
- 

**4.2. APX-Hardness**

In this section, we prove that MIS, MDS, MVC remain APX-hard even on power-law graphs.

**Theorem 4.2.** *MDS is APX-hard on power-law graphs.*

**Proof.** According to Theorem 4.1, we use the cycle-based embedding technique to show  $\mathcal{L}$ -reduction from MDS on any  $d$ -bounded graph  $G_d$  to MDS on a power-law graph  $G_{(\alpha, \beta)}$  since MDS is proven APX-hard on  $d$ -bounded graphs [22].

Letting  $\phi$  be a feasible solution on  $G_d$ , we can construct MDS in  $G'$  such that MDS on a  $K_2$  is 1,  $n/4$  on a  $\vec{d}$ -regular cycle and  $n/3$  on a cycle and a  $\vec{\kappa}$ -branch- $\vec{d}$ -cycle. Therefore, for a solution  $\phi$  on  $G_d$ , we have a solution  $\varphi$  on  $G_{(\alpha, \beta)}$  to be  $\varphi = \phi + n_1/2 + n_2/3 + n_3/4$ , where  $n_1, n_2$  and  $n_3$  corresponds to  $\tau(1), \tau(2) \cup L$  and all leftover vertices. Hence, we have  $OPT(\varphi) = OPT(\phi) + n_1/2 + n_2/3 + n_3/4$ .

On one hand, for a  $d$ -bounded graph with vertices  $n$ , the optimal MDS is lower bounded by  $n/(d+1)$ . Thus, we know

$$\begin{aligned} OPT(\varphi) &= OPT(\phi) + n_1/2 + n_2/3 + n_3/4 \\ &\leq OPT(\phi) + (N - n)/2 \leq OPT(\phi) + (\zeta(\beta)d^\beta - 1)n/2 \\ &\leq OPT(\phi) + (\zeta(\beta)d^\beta - 1)(d+1)OPT(\phi)/2 = \left[1 + (\zeta(\beta)d^\beta - 1)(d+1)/2\right] OPT(\phi) \end{aligned}$$

where  $N$  is the number of vertices in  $G_{(\alpha, \beta)}$ .

On the other hand, with  $|OPT(\phi) - \phi| = |OPT(\varphi) - \varphi|$ , we proved the  $\mathcal{L}$ -reduction with  $c_1 = 1 + (\zeta(\beta)d^\beta - 1)(d+1)/2$  and  $c_2 = 1$ .  $\square$

**Theorem 4.3.** *MVC is APX-hard on power-law graphs.*

**Proof.** In this proof, we show  $\mathcal{L}$ -reduction from MVC on  $d$ -bounded graph  $G_d$  to MVC on power-law graph  $G_{(\alpha, \beta)}$  using cycle-based embedding technique.

Let  $\phi$  be a feasible solution on  $G_d$ . We construct the solution  $\varphi \leq \phi + (N - n)$  since the optimal solution of MVC is  $n/2$  on  $K_2$ , cycle,  $\vec{d}$ -regular cycle and  $n$  on  $\vec{\kappa}$ -branch- $\vec{d}$ -cycle. Therefore, since the optimal MVC on a  $d$ -bounded graph is lower bounded by  $n/(d+1)$ , we have

$$OPT(\varphi) \leq \left[1 + (\zeta(\beta)d^\beta - 1)(d+1)\right] OPT(\phi)$$

On the other hand, with  $|OPT(\phi) - \phi| = |OPT(\varphi) - \varphi|$ , we proved the  $\mathcal{L}$ -reduction with  $c_1 = 1 + (\zeta(\beta)d^\beta - 1)(d+1)$  and  $c_2 = 1$ .  $\square$

**Corollary 4.1.** *MIS is APX-hard on power-law graphs.*

**4.3. Inapproximability factors**

In this section, we show the inapproximability factors on MIS, MVC and MDS on power-law graphs respectively using the results in Section 2.4.

**Theorem 4.4.** *For any  $\varepsilon > 0$ , there is no  $1 + \frac{1}{140(2\zeta(\beta)3^\beta - 1)} - \varepsilon$  approximation algorithm for Maximum Independent Set on power-law graphs.*

**Proof.** In this proof, we construct the power-law graph  $G_{(\alpha, \beta)}$  based on cycle-based embedding technique in Theorem 4.1 from  $d$ -bounded graph  $G_d$ . Let  $\phi$  and  $\varphi$  be feasible solutions of MIS on  $G_d$  and  $G_{(\alpha, \beta)}$ . Then  $OPT(\varphi)$  composed of  $OPT(\phi)$ , clique  $K_2$ , cycle,  $\vec{d}$ -regular cycle and  $\vec{\kappa}$ -branch- $\vec{d}$ -cycles are all exactly half number of vertices. Hence, we have  $OPT(\varphi) = OPT(\phi) + (N - n)/2$  where  $n$  and  $N$  is the number of vertices in  $G_d$  and  $G_{(\alpha, \beta)}$  respectively. Since  $OPT(\phi) \geq n/(d+1)$  on  $d$ -bounded graphs for MIS and  $N \leq \zeta(\beta)d^\beta n$ , we further have  $\mathfrak{C} = 1 + \frac{\zeta(\beta)d^\beta - 1}{2(d+1)}$  from

$$\begin{aligned} OPT(\varphi) &= OPT(\phi) + \frac{N - n}{2} \leq OPT(\phi) + \frac{(\zeta(\beta)d^\beta - 1)}{2}n \\ &\leq OPT(\phi) + \frac{(\zeta(\beta)d^\beta - 1)(d+1)}{2}OPT(\phi) = \left(1 + \frac{(\zeta(\beta)d^\beta - 1)(d+1)}{2}\right) OPT(\phi) \end{aligned}$$



According to  $\epsilon = \frac{140}{139} - \epsilon'$  for any  $\epsilon' > 0$  on 3-bounded graphs, then the inapproximability factor can be derived from inapproximability optimal substructure framework as

$$\delta > \frac{\epsilon \mathfrak{C}}{(\mathfrak{C} - 1)\epsilon + 1} > 1 + \frac{1}{140\mathfrak{C}} - \epsilon = 1 + \frac{1}{140(2\zeta(\beta)3^\beta - 1)} - \epsilon$$

where the last step follows from  $d = 3$ .  $\square$

**Theorem 4.5.** *There is no  $1 + \frac{1}{390(2\zeta(\beta)3^\beta - 1)}$  approximation algorithm for Minimum Dominating Set on power-law graphs.*

**Proof.** In this proof, we construct the power-law graph  $G_{(\alpha, \beta)}$  based on cycle-based embedding technique in Theorem 4.1 from  $d$ -bounded graph  $G_d$ . Let  $\phi$  and  $\varphi$  be feasible solutions of MDS on  $G_d$  and  $G_{(\alpha, \beta)}$ . The optimal MDS on  $OPT(\phi)$ , clique  $K_2$ , cycle,  $\vec{d}$ -regular cycle and  $\vec{\kappa}$ -branch- $\vec{d}$ -cycles are  $n/2$ ,  $n/4$  and  $n/3$  respectively. Let  $\phi$  and  $\varphi$  be feasible solutions of MDS on  $G_d$  and  $G_{(\alpha, \beta)}$ . Then we have  $\mathfrak{C} = 1 + \frac{(\zeta(\beta)d^\beta - 1)(d+1)}{2}$  similar as the proof in Theorem 4.4.

According to  $\epsilon = \frac{391}{390}$  in 3-bounded graphs, then the inapproximability factor can be derived from inapproximability optimal substructure framework as

$$\delta > 1 + \frac{\epsilon - 1}{\mathfrak{C}} = 1 + \frac{1}{390(2\zeta(\beta)3^\beta - 1)}$$

where the last step follows from  $d = 3$ .  $\square$

**Theorem 4.6.** *MVC is hard to be approximated within  $1 + \frac{2(1 - (2 + o_c(1)) \frac{\log \log c}{\log c})}{\left(\zeta(\beta)c^\beta + c^{\frac{1}{\beta}}\right)(c+1)}$  on power-law graphs under unique games conjecture.*

**Proof.** By constructing the power-law graph  $G_{(\alpha, \beta)}$  based on cycle-based embedding technique in Theorem 4.1 from  $d$ -bounded graph  $G_d$ , The optimal MVC on clique  $K_2$ , cycle,  $\vec{d}$ -regular cycle consists of half number of vertices while the optimal MVC on  $\vec{\kappa}$ -branch- $\vec{d}$ -cycles consists of all vertices. Thus, we have  $\mathfrak{C} = 1 + \frac{\left(\zeta(\beta)d^\beta - 1 + d^{\frac{1}{\beta}}\right)(d+1)}{2}$  since

$$OPT(\varphi) \leq OPT(\phi) + \frac{N - n - \Delta}{2} + \Delta \leq OPT(\phi) + \frac{(\zeta(\beta)d^\beta - 1)n + n^{\frac{1}{\beta}}d}{2} \quad (4.1)$$

$$= OPT(\phi) + \frac{\left(\zeta(\beta)d^\beta - 1 + \frac{d}{n^{1-\frac{1}{\beta}}}\right)n}{2} \quad (4.2)$$

$$\leq OPT(\phi) + \frac{\left(\zeta(\beta)d^\beta - 1 + \frac{d}{(d+1)^{1-\frac{1}{\beta}}}\right)(d+1)}{2} OPT(\phi) \quad (4.3)$$

$$\leq \left(1 + \frac{\left(\zeta(\beta)d^\beta - 1 + d^{\frac{1}{\beta}}\right)(d+1)}{2}\right) OPT(\phi) \quad (4.4)$$

where  $\phi$  and  $\varphi$  are feasible solutions of MVC on  $G_d$  and  $G_{(\alpha, \beta)}$ ,  $\Delta$  is the maximum degree in  $G_{(\alpha, \beta)}$ . The inequality (4.1) holds since there are at most  $\Delta$  vertices in  $\vec{\kappa}$ -branch- $\vec{d}$ -cycle, i.e.  $\Delta = e^{\alpha/\beta} \leq n^{1/\beta}d$ ; (4.3) holds since there are at least  $d+1$  vertices in a  $d$ -bounded graph and the optimal MVC in a  $d$ -bounded graph is at least  $n/(d+1)$ .

According to  $\epsilon = 2 - (2 + o_d(1)) \log \log d / \log d$ , then the inapproximability factor can be derived from inapproximability optimal substructure framework as

$$\delta > 1 + \frac{\epsilon - 1}{\mathfrak{C}} \geq 1 + \frac{2\left(1 - (2 + o_c(1)) \frac{\log \log c}{\log c}\right)}{\left(\zeta(\beta)c^\beta + c^{\frac{1}{\beta}}\right)(c+1)}$$

where  $c$  is the smallest  $d$  satisfying the condition in [15]. The last inequality holds since function  $f(x) = (1 - (2 + o_x(1)) \log \log x / \log x) / g(x)(x+1)$  is monotonously decreasing when  $f(x) > 0$  for all  $x > 0$  when  $g(x)$  is monotonously increasing.  $\square$

**Theorem 4.7.**  *$\rho$ -PDS is hard to be approximated into  $2 - (2 + o_d(1)) \frac{\log \log d}{\log d}$  on  $d$ -bounded graphs under unique games conjecture.*

**Proof.** In this proof, we show the gap-preserving reduction from MVC on  $(d/\rho)$ -bounded graph  $G = (V, E)$  to  $\rho$ -PDS on  $d$ -bounded graph  $G' = (V', E')$ . w.l.o.g., we assume that  $d$  and  $d/\rho$  are integers. We construct a graph  $G' = (V', E')$  by adding

new vertices and edges to  $G$  as follows. For each edge  $(v_i, v_j) \in E$ , create  $k$  new vertices  $v_{ij}^1, \dots, v_{ij}^k$  where  $1 \leq k \leq \lfloor 1/\rho \rfloor$  and  $\rho \leq 1/2$ . Then we add  $2k$  new edges  $(v_{ij}^l, v_i)$  and  $(v_{ij}^l, v_j)$  for all  $l \in [1, k]$  as shown in Fig. 2. Clearly,  $G' = (V', E')$  is a  $d$ -bounded graph.

Let  $\phi$  and  $\varphi$  be feasible solutions to MVC on  $G$  and  $G'$  respectively. We claim that  $OPT(\phi) = OPT(\varphi)$ .

On one hand, if  $S = \{v_1, v_2, \dots, v_j\} \in V$  is the minimum vertex cover on  $G$ . Then  $\{v_1, v_2, \dots, v_j\}$  is a  $\rho$ -PDS on  $G'$  because each vertex in  $V$  has  $\rho$  of all neighbors in MVC and every new vertex in  $V' \setminus V$  has at least one of two neighbors in MVC. Thus  $OPT(\phi) \geq OPT(\varphi)$ . On the other hand, we can prove that  $OPT(\varphi)$  does not contain new vertices, that is,  $V' \setminus V$ . Consider a vertex  $v_i \in V$ , if  $v_i \in OPT(\varphi)$ , the new vertices  $v_{ij}^l$  for all  $v_j \in N(v_i)$  and all  $l \in [1, k]$  are not needed to be selected. If  $v_i \notin OPT(\varphi)$ , it has to be dominated by  $\rho$  proportion of its all neighbors. That is, for each edge  $(v_i, v_j)$  incident to  $v_i$ , either  $v_j$  or all  $v_{ij}^l$  have to be selected since every  $v_{ij}^l$  has to be either selected or dominated. If all  $v_{ij}^l$  are selected in  $OPT(\varphi)$  for some edge  $(v_i, v_j)$ ,  $v_j$  is still not dominated by enough vertices if there are some more edges incident to  $v_j$  and the number of vertices  $v_{ij}^l$  is great than 1, that is,  $\lfloor 1/\rho \rfloor \geq 1$ . In this case,  $v_j$  will be selected to dominate all  $v_{ij}^l$ . Thus,  $OPT(\varphi)$  does not contain new vertices. Since the vertices in  $V$  selected is a solution to  $\rho$ -MDS, that is, for each vertex  $v_i$  in graph  $G$ ,  $v_i$  will be selected or at least the number of neighbors of  $v_i$  will be selected. Therefore, the vertices in  $OPT(\varphi)$  consist of a vertex cover in  $G$ . Thus  $OPT(\phi) \leq OPT(\varphi)$ . Then we show the completeness and soundness as follows.

- If  $OPT(\phi) = m \Rightarrow OPT(\varphi) = m$
- If  $OPT(\phi) > \left(2 - (2 + o_d(1)) \frac{\log \log(d/2)}{\log(d/2)}\right) m \Rightarrow OPT(\varphi) > \left(2 - (2 + o_d(1)) \frac{\log \log d}{\log d}\right) m$

$$OPT(\varphi) > \left(2 - (2 + o_d(1)) \frac{\log \log(d/\rho)}{\log(d/\rho)}\right) m > \left(2 - (2 + o_d(1)) \frac{\log \log d}{\log d}\right) m$$

since the function  $f(x) = 2 - \log \log x / \log x$  is monotonously increasing for any  $x > 0$ .  $\square$

**Theorem 4.8.**  $\rho$ -PDS is hard to be approximated into  $1 + \frac{2(1 - (2 + o_c(1)) \frac{\log \log c}{\log c})}{2 + (\zeta(\beta)c^\beta - 1)(c + 1)}$  on power-law graphs under unique games conjecture.

**Proof.** By constructing the power-law graph  $G_{(\alpha, \beta)}$  based on cycle-based embedding technique in Theorem 4.1 from  $d$ -bounded graph  $G_d$ , and according to the optimal MVC on  $OPT(\phi)$ , clique  $K_2$ , cycle,  $\vec{d}$ -regular cycle and  $\vec{\kappa}$ -branch- $\vec{d}$ -cycles, we have  $\mathfrak{C} = 1 + \frac{(\zeta(\beta)d^\beta - 1)(d + 1)}{2}$  from

$$\begin{aligned} OPT(\varphi) &= OPT(\phi) + n_1/2 + f(\rho)n_2 + g(\rho)n_3 \\ &\leq OPT(\phi) + \frac{N - n}{2} \leq \left(1 + \frac{(\zeta(\beta)d^\beta - 1)(d + 1)}{2}\right) OPT(\phi) \end{aligned}$$

where  $f(\rho) = \begin{cases} \frac{1}{4}, & \rho \leq \frac{1}{3} \\ \frac{1}{3}, & \frac{1}{3} < \rho \leq \frac{1}{2} \end{cases}$ ,  $g(\rho) = \frac{1}{3}$  for all  $\rho \leq \frac{1}{2}$  and  $\phi, \varphi$  be feasible solutions of MVC on  $G_d$  and  $G_{(\alpha, \beta)}$ .  $n_1, n_2$  and  $n_3$  are correspondent to the number of vertices in cliques  $K_2$ , cycle,  $\vec{d}$ -regular cycle and  $\vec{\kappa}$ -branch- $\vec{d}$ -cycle.

According to  $\epsilon = 2 - (2 + o_d(1)) \log \log d / \log d$ , then the inapproximability factor can be derived from inapproximability optimal substructure framework as

$$\delta > 1 + \frac{\epsilon - 1}{\mathfrak{C}} \geq 1 + \frac{2(1 - (2 + o_c(1)) \frac{\log \log c}{\log c})}{2 + (\zeta(\beta)c^\beta - 1)(c + 1)}$$

where  $c$  is the smallest  $d$  satisfying the condition in [15]. The last inequality holds since function  $f(x) = (1 - (2 + o_x(1)) \log \log x / \log x) / g(x)(x + 1)$  is monotonously decreasing when  $f(x) > 0$  for all  $x > 0$  when  $g(x)$  is monotonously increasing.  $\square$

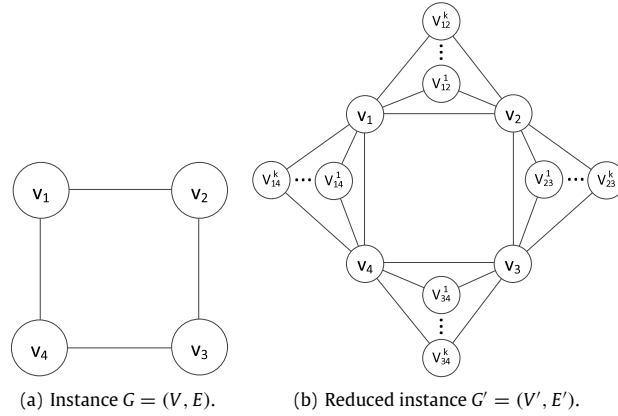
## 5. More inapproximability results on simple power-law graphs

### 5.1. General graphic embedding technique

In this section, we introduce a general graphic embedding technique to embed a  $d$  bounded graph into a simple power-law graph. Before presenting the embedding technique, we first show that a graph can be constructed in polynomial time from a class of integer sequences.

**Lemma 5.1.** Given a sequence of integers  $D = \langle d_1, d_2, \dots, d_n \rangle$  which is non-increasing, continuous and the number of elements is at least twice as the largest element in  $D$ , i.e.  $n \geq 2d_1$ , it is possible to construct a simple graph  $G$  whose  $d$ -degree sequence is  $D$  in polynomial time  $O(n^2 \log n)$ .



Fig. 2. The reduction from MVC to  $\rho$ -MDS.

**Proof.** Starting with a set of individual vertices  $S$  of degree 0 and  $|S| = n$ , we iteratively connect vertices together to increase their degrees up to given degree sequence. In each step, the leftover vertex of highest degree is connected to other vertices one by one in the decreasing order of their degrees. Then the sequence  $D$  will be resorted and all zero elements will be removed. The algorithm stops until  $D$  is empty. The whole algorithm is shown as follows (Algorithm 2).

---

**Algorithm 2:** Graphic Sequence Construction Algorithm

---

**Input :**  $d$ -degree sequence  $D = \langle d_1, d_2, \dots, d_n \rangle$  where  $d_1 \geq d_2 \geq \dots \geq d_n$

**Output:** Graph  $H$

```

1 while  $D \neq \emptyset$  do
2   Connect vertex of  $d_1$  to vertices of  $d_2, d_3, \dots, d_{d_1+1}$ ;
3    $d_1 \leftarrow 0$ ;
4   for  $i = 2$  to  $d_1 + 1$  do
5      $d_i \leftarrow d_i - 1$ ;
6   end
7   Sort  $D$  in non-increasing order;
8   Remove all zero elements in  $D$ ;
9 end

```

---

After each while loop, the new degree sequence, called  $D'$ , is still continuous and its number of elements is at least as twice as its maximum element. To show this, we consider three cases: (1) If the maximum degree in  $D'$  remains the same, there are at least  $d_1 + 2$  vertices in  $D$ . Since  $D$  is continuous, the number of elements in  $D$  is at least  $d_1 + 2 + d_1 - 1$ , that is,  $2d_1 + 1$ . Therefore, the number of elements in  $D'$  is  $2d_1$ , i.e.  $n \geq 2d_1$  still holds. (2) If the maximum degree in  $D'$  is decreased by 1, there are at least 2 elements of degree  $d_1$  in  $D$ . Thus, at most one element in  $D$  will become 0. Then we have  $n \geq 2d_1 - 2 = 2(d_1 - 1)$ . (3) If the maximum degree in  $D'$  is decreased by 2, there are at most two elements in  $D$  becoming 0. Thus,  $n \geq 2d_1 - 3 > 2(d_1 - 2)$ .

The time complexity of the algorithm is  $O(n^2 \log n)$  since there are at most  $n$  iterations and each iteration takes at most  $O(n \log n)$  to sort the new sequence  $D$ .  $\square$

**Theorem 5.1** (Graphic Embedding Technique). Any  $d$ -bounded graph  $G_d$  can be embedded into a simple power-law graph  $G_{(\alpha, \beta)}$  with  $\beta > 1$  in polynomial time such that  $G_d$  is a maximal component and the number of vertices in  $G_{(\alpha, \beta)}$  can be polynomially bounded by the number of vertices in  $G_d$ .

**Proof.** Given a  $d$ -bounded degree graph  $G_d = (V, E)$  and  $\beta > 1$ , we construct a power-law graph  $G_{(\alpha, \beta)}$  of exponential factor  $\beta$  which includes  $G_d$  as a set of maximal components. The construction is shown as Algorithm 3.

---

**Algorithm 3:** Graphic Embedding Algorithm

---

```

1  $\alpha \leftarrow \max\{\frac{\beta}{\beta-1}(\ln 4 + \beta \ln d), \ln 2 + \ln n + \beta \ln d\}$  and corresponding  $G_{(\alpha, \beta)}$ ;
2  $D$  be the  $d$ -degree sequence of  $G_{(\alpha, \beta)} \setminus G_d$ ;
3 Construct  $G_{(\alpha, \beta)} \setminus G_d$  using Algorithm 2.

```

---

According to the [Lemma 5.1](#), the above construction is valid and finishes in polynomial time. Then we show that  $N$  is upper bounded by  $\zeta(\beta)2d^\beta n$ , where  $n$  and  $N$  are the number of vertices in  $G_d$  and  $G_{\alpha,\beta}$  respectively. From the construction, we know either

$$\alpha \geq \frac{\beta}{\beta-1}(\ln 4 + \beta \ln d) \Rightarrow \alpha \geq \ln 4 + \beta \ln d + \alpha/\beta \Rightarrow \frac{e^\alpha}{d^\beta} \geq 4e^{\frac{\alpha}{\beta}}$$

or

$$\alpha \geq \ln 2 + \ln n + \beta \ln d \Rightarrow \frac{e^\alpha}{d^\beta} \geq 2n$$

Therefore,  $\frac{e^\alpha}{d^\beta} \geq 2e^{\frac{\alpha}{\beta}} + n$ . Note that  $\left\lfloor \frac{e^\alpha}{d^\beta} \right\rfloor$  is the number of vertices of degree  $d$ . In addition,  $G$  has at most  $n$  vertices of degree  $d$ , so  $D$  is continuous degree sequence and has the number of vertices at least as twice as the maximum degree.

In addition, when  $n$  is large enough, we have  $\alpha = \ln 2 + \ln n + \beta \ln d$ . Hence, the number of vertices  $N$  in  $G_{\alpha,\beta}$  is bound as  $N \leq \zeta(\beta)e^\alpha = 2\zeta(\beta)d^\beta n$ , i.e. the number of vertices of  $G_{\alpha,\beta}$  is polynomial bounded by the number of vertices in  $G_d$ .  $\square$

## 5.2. Inapproximability of MIS, MVC and MDS

**Theorem 5.2.** For any  $\varepsilon > 0$ , it is NP-hard to approximate Maximum Independent Set within  $1 + \frac{1}{1120\zeta(\beta)3^\beta} - \varepsilon$  on simple power-law graphs.

**Proof.** In this proof, we construct the simple power-law graph  $G_{(\alpha,\beta)}$  based on graphic embedding technique in [Theorem 5.1](#) from  $d$ -bounded graph  $G_d$ . Let  $\phi$  and  $\varphi$  be feasible solutions of MIS on  $G_d$  and  $G_{(\alpha,\beta)}$ . Since  $\text{OPT}(\phi) \geq n/(d+1)$  on  $d$ -bounded graphs and  $N \leq 2\zeta(\beta)d^\beta n$ , we further have  $\mathfrak{C} = 2\zeta(\beta)d^\beta(d+1)$  from

$$\text{OPT}(\varphi) \leq N \leq 2\zeta(\beta)d^\beta n \leq 2\zeta(\beta)d^\beta(d+1)\text{OPT}(\phi)$$

According to  $\epsilon = \frac{140}{139} - \epsilon'$  for any  $\epsilon' > 0$  on 3-bounded graphs, then the inapproximability factor can be derived from inapproximability optimal substructure framework as

$$\delta > \frac{\epsilon \mathfrak{C}}{(\mathfrak{C}-1)\epsilon + 1} = 1 + \frac{1}{140\mathfrak{C}-1} - \epsilon > 1 + \frac{1}{1120\zeta(\beta)3^\beta} - \epsilon. \quad \square$$

**Theorem 5.3.** It is NP-hard to approximate Minimum Dominating Set within  $1 + \frac{1}{3120\zeta(\beta)3^\beta}$  on power-law graphs.

**Proof.** From the proof of [Theorem 5.2](#), we have  $\mathfrak{C} = 2\zeta(\beta)d^\beta(d+1)$ . Then according to  $\epsilon = \frac{391}{390}$  on 3-bounded graphs, we have

$$\delta > 1 + \frac{\epsilon - 1}{\mathfrak{C}} \geq 1 + \frac{1}{3120\zeta(\beta)3^\beta}. \quad \square$$

**Theorem 5.4.** There is no  $1 + \frac{2-(2+o_c(1))\frac{\log \log c}{\log c}}{2\zeta(\beta)c^\beta(c+1)}$  approximation algorithm of Minimum Vertex Cover on power-law graphs under unique games conjecture.

**Proof.** Similar as the proof of [Theorem 5.3](#), we have  $\mathfrak{C} = 2\zeta(\beta)d^\beta(d+1)$ . Then according to  $\epsilon = 2 - (2 + o_d(1))\log \log d / \log d$ , then the inapproximability factor can be derived from inapproximability optimal substructure framework as

$$\delta > 1 + \frac{\epsilon - 1}{\mathfrak{C}} \geq 1 + \frac{2 - (2 + o_c(1))\frac{\log \log c}{\log c}}{2\zeta(\beta)c^\beta(c+1)}$$

where  $c$  is the smallest  $d$  satisfying the condition in [15].  $\square$

**Theorem 5.5.** There is no  $1 + \frac{2-(2+o_c(1))\frac{\log \log c}{\log c}}{2\zeta(\beta)c^\beta(c+1)}$  approximation algorithm for Minimum Positive Dominating Set on power-law graphs.

**Proof.** Similar as [Theorem 5.5](#), the proof follows from [Theorem 4.7](#).  $\square$

### 5.3. Maximum clique, minimum coloring

**Lemma 5.2** (Ferrante et al. [1]). Let  $G = (V, E)$  be a simple graph with  $n$  vertices and  $\beta \geq 1$ . Let  $\alpha \geq \max\{4\beta, \beta \log n + \log(n+1)\}$ . Then,  $G_2 = G \setminus G_1$  is a bipartite graph.

**Lemma 5.3.** Given a function  $f(x)$  ( $x \in \mathbb{Z}, f(x) \in \mathbb{Z}^+$ ) monotonously decreases, then  $\sum_x f(x) \leq \int_x f(x)$ .

**Corollary 5.1.**  $e^\alpha \sum_{i=1}^{e^{\alpha/\beta}} \left(\frac{1}{i}\right)^\beta < (e^\alpha - e^{\alpha/\beta})/(\beta - 1)$ .

**Theorem 5.6.** Maximum Clique cannot be approximated within  $O(n^{1/(\beta+1)-\epsilon})$  on large power-law graphs with  $\beta > 1$  and  $n > 54$  for any  $\epsilon > 0$  unless  $NP = ZPP$ .

**Proof.** In [1], the authors proved the hardness of Maximum Clique problem on power-law graphs. Here we use the same construction. According to Lemma 5.2,  $G_2 = G \setminus G_1$  is a bipartite graph when  $\alpha \geq \max\{4\beta, \beta \log n + \log(n+1)\}$  for any  $\beta \geq 1$ . Let  $\phi$  be a solution on general graph  $G$  and  $\varphi$  be a solution on power-law graph  $G_2$ . We show the completeness and soundness.

- If  $OPT(\phi) = m \Rightarrow OPT(\varphi) = m$   
If  $OPT(\phi) \leq 2$  on graph  $G$ , we can solve clique problem in polynomial time by iterating the edges and their endpoints one by one. However,  $G$  is not a general graph in this case. w.l.o.g, assuming  $OPT(\phi) > 2$ , then  $OPT(\varphi) = OPT(\phi) > 2$  since the maximum clique on bipartite graph is 2.
- If  $OPT(\phi) \leq m/n^{1-\epsilon} \Rightarrow OPT(\varphi) < O\left(1/(N^{1/(\beta+1)-\epsilon'})\right) m$   
In this case, we consider the case that  $4\beta < \beta \log n + \log(n+1)$ , that is,  $n > 54$ . According to Lemma 5.2, let  $\alpha = \beta \log n + \log(n+1)$ . From Corollary 5.1, we have

$$N = e^\alpha \sum_{i=1}^{\Delta} \left(\frac{1}{i}\right)^\beta < \frac{e^\alpha - e^{\alpha/\beta}}{\beta - 1} = \frac{n^\beta(n+1) - n(n+1)^{1/\beta}}{\beta - 1} < \frac{2n^{\beta+1} - n}{\beta - 1}$$

Therefore,  $OPT(\varphi) = OPT(\phi) \leq m/n^{1-\epsilon} < O\left(m/(N^{1/(\beta+1)-\epsilon'})\right)$ .  $\square$

**Corollary 5.2.** Minimum Coloring problem cannot be approximated within  $O(n^{1/(\beta+1)-\epsilon})$  on large power-law graphs with  $\beta > 1$  and  $n > 54$  for any  $\epsilon > 0$  unless  $NP = ZPP$ .

## 6. Relationship between $\beta$ and approximation hardness

As shown in previous sections, many hardness and inapproximability results are dependent on  $\beta$ . In this section, we analyze the hardness of some optimal substructure problems based on  $\beta$  by showing that trivial greedy algorithms can achieve constant guarantee factors for MIS and MDS.

**Lemma 6.1.** When  $\beta > 2$ , the size of MDS of a power-law graph is greater than  $Cn$  where  $n$  is the number of vertices,  $C$  is some constant only dependent on  $\beta$ .

**Proof.** Let  $S = (v_1, v_2, \dots, v_t)$  of degrees  $d_1, d_2, \dots, d_t$  be the MDS of power-law graph  $G_{(\alpha, \beta)}$ . Observing that the total degrees of vertices in dominating set must be at least the number of vertices outside the dominating set, we have  $\sum_{i=1}^{i=t} d_i \geq |V \setminus S|$ . With a given total degree, a set of vertices has minimum size when it includes the vertices of highest degrees. Since the function  $\zeta(\beta - 1) = \sum_{i=1}^{\infty} \frac{1}{i^{\beta-1}}$  converges when  $\beta > 2$ , there exists a constant  $t_0 = t_0(\beta)$  such that

$$\sum_{i=t_0}^{\Delta} i \left\lfloor \frac{e^\alpha}{i^\beta} \right\rfloor \geq \sum_{i=1}^{t_0} \left\lfloor \frac{e^\alpha}{i^\beta} \right\rfloor$$

where  $\alpha$  is any large enough constant. Thus the size of MDS is at least

$$\sum_{i=t_0}^{\Delta} \left\lfloor \frac{e^\alpha}{i^\beta} \right\rfloor \approx \left( \zeta(\beta) - \sum_{i=1}^{t_0-1} \frac{1}{i^\beta} \right) e^\alpha \approx C|V|$$

where  $C = (\zeta(\beta) - \sum_{i=1}^{t_0} \frac{1}{i^\beta})/(\zeta(\beta))$ .  $\square$

Consider the greedy algorithm which selects from the vertices of the highest degree to the lowest. In the worst case, it selects all vertices with degree greater than 1 and a half of vertices with degree 1 to form a dominating set. The approximation factor of this simple algorithm is a constant.

**Corollary 6.1.** Given a power-law graph with  $\beta > 2$ , the greedy algorithm that selects vertices in decreasing order of degrees provides a dominating set of size at most  $\sum_{i=2}^{\Delta} \lfloor e^{\alpha}/i^{\beta} \rfloor + \frac{1}{2}e^{\alpha} \approx (\zeta(\beta) - 1/2)e^{\alpha}$ . Thus the approximation ratio is  $(\zeta(\beta) - \frac{1}{2})/(\zeta(\beta) - \sum_{i=1}^{i_0} 1/i^{\beta})$ .

Let us consider another maximization problem MIS, we propose a greedy algorithm Power-law-Greedy-MIS as follows. We sort the vertices in non-increasing order of degrees and start checking from the vertex of lowest degree. If the vertex is not adjacent to any selected vertex, it is selected. The set of selected vertices forms an independent set with the size at least a half the number of vertices of degree 1 which is  $e^{\alpha}/2$ . The size of MIS is at most a half of number of vertices. Thus, the following lemma holds.

**Lemma 6.2.** Power-law-Greedy-MIS has factor  $1/(2\zeta(\beta))$  on power-law graphs with  $\beta > 1$ .

## 7. Conclusion

In this paper, we analyzed the approximation hardness and inapproximability of optimal substructure problems on power-law graphs. These problems are only illustrated in the literature not be able to approximated into some constant factors on both general and simple power-law graphs although they remain APX-hard. However, we also notice that the gap between inapproximability factor and the simple constant approximation ratio of these problems is still not small enough and the hardness on power-law graph is weaker than that on degree bounded graphs. A question is raised: Is there any efficient reduction which is not from bounded graph will improve the hardness results on power-law graphs? Can we get stronger hardness results based on some specific power-law models? For example, if the number of vertices only follow power-law distribution when degree is larger than some constant  $i_0$ , we can reduce from graph of degree bounded by  $i_0$  and get better results.

On the contrary, we also show that MAX CLIQUE and GRAPH COLORING are still very hard to be approximated since the optimal solutions to these problems are dependent on the structure of local graph component rather than global graph. In other words, the power-law distribution in degree sequence does not help much for such optimization problems without the property of optimal substructure.

## Acknowledgement

This work is partially supported by NSF Career Award # 0953284, DTRA, Young Investigator Award, Basic Research Program # HDTRA1-09-1-0061 and DTRA # HDTRA1-08-10.

## Appendix. Embedding construction with $\beta < 1$

Ferrante et al. [1] proved the NP-hardness of MIS, MDS, and MVC where  $\beta < 1$  based on Lemma A.1 which is invalid. A counter-example is as follows. Let  $D_1 = \langle 3, 2, 2, 1 \rangle$  and  $D_2 = \langle 7, 6, 5, 4, 3, 2, 2, 1 \rangle$  then  $D_1$  is eligible and  $Y_1 = \langle 1, 2, 1 \rangle$ ,  $Y_2 = \langle 1, 2, 1, 1, 1, 1, 1 \rangle$  but  $D_2$  is NOT eligible with  $f_{D_2}(4) < 0$ . In this appendix, we present an alternative lemma to prove the hardness of these problems on power-law graphs with  $\beta < 1$ .

**Definition A.1** (*d-Degree Sequence*). Given a graph  $G = (V, E)$ , the  $d$ -degree sequence of  $G$  is a sequence  $D = \langle d_1, d_2, \dots, d_n \rangle$  of vertex degrees in non-increasing order.

**Definition A.2** (*y-Degree Sequence*). Given a graph  $G = (V, E)$ , the  $y$ -degree sequence of  $G$  is a sequence  $Y = \langle y_1, y_2, \dots, y_m \rangle$  where  $m$  is the maximum degree of  $G$  and  $y_i = |\{u \in V \mid \deg(u) = i\}|$ .

**Definition A.3** (*Eligible Sequences*). A sequence of integers  $S = \langle s_1, \dots, s_n \rangle$  is eligible if  $s_1 \geq s_2 \geq \dots \geq s_n$  and  $f_S(k) \geq 0$  for all  $k \in [n]$ , where

$$f_S(k) = k(k-1) + \sum_{i=k+1}^n \min\{k, s_i\} - \sum_{i=1}^k s_i$$

**Lemma A.1** (*Invalid Lemma in [1]*). Let  $Y_1$  and  $Y_2$  be two  $y$ -degree sequences with  $m_1$  and  $m_2$  elements respectively such that (1)  $Y_1(i) \leq Y_2(i)$ ,  $\forall i \leq m_1$ , and (2) two corresponding  $d$ -degree sequences  $D_1$  and  $D_2$  are contiguous. If  $D_1$  is eligible then  $D_2$  is eligible.

Erdős and Gallai [23] showed that an integer sequence is graphic -  $d$ -degree sequence of an graph, if and only if it is eligible and the total of all elements is even. Then Havel and Hakimi [24] gave an algorithm to construct a simple graph from a degree sequence.

**Lemma A.2** ([24]). A sequence of integers  $D = \langle d_1, \dots, d_n \rangle$  is graphic if and only if it is non-increasing and the sequence of values  $D' = \langle d_2 - 1, d_3 - 1, \dots, d_{d_1+1} - 1, d_{d_1+2}, \dots, d_n \rangle$  sorted in non-increasing order is graphic.

We now prove the following lemma, which can substitute Lemma A.1 for the NP-hardness proof in [1].

**Lemma A.3.** *Given an undirected graph  $G = (V, E)$ ,  $0 < \beta < 1$ , there exists polynomial time algorithm to construct a power-law graph  $G' = (V', E')$  of exponential factor  $\beta$  such that  $G$  is a set of maximal components of  $G'$ .*

**Proof.** To construct  $G'$ , we choose  $\alpha = \max\{\beta \ln(n-1) + \ln(n+2), 3 \ln 2\}$ . Then  $\lfloor e^\alpha / ((n-1)^\beta) \rfloor > n+2$ , i.e. there are at least 2 vertices of degree  $d$  in  $G' \setminus G$  if there are at least 2 vertices of degree  $d$  in  $G$ . According to the definition, the total degrees of all vertices in  $G'$  and  $G$  are even. Therefore, the lemma will follow if we prove that the degree sequence  $D$  of  $G' \setminus G$  is eligible.

In  $D$ , the maximum degree is  $\lfloor e^{\alpha/\beta} \rfloor$ . There is only one vertex of degree  $i$  if  $1 \leq e^\alpha / i^\beta < 2$ , i.e.  $e^{\alpha/\beta} \geq i > (e^\alpha / 2)^{1/\beta}$ .

Let us consider  $f_D(k)$  in two cases:

1. **Case 1:**  $k \leq \lfloor e^{\alpha/\beta} / 2 \rfloor$

$$\begin{aligned} f_D(k) &= k(k-1) + \sum_{i=k+1}^n \min\{k, d_i\} - \sum_{i=1}^k d_i \\ &> k(k-1) + \sum_{i=k}^{T-k} k + \sum_{i=B}^{k-1} i + \sum_{i=1}^{B-1} 2 - \sum_{i=1}^k (T-k+1) \\ &= k(T-k) + (k-B)(k-1+B)/2 + B(B-1) - k(2T-k+1)/2 \\ &= (B^2 - B)/2 - k \end{aligned}$$

where  $T = \lfloor e^{\alpha/\beta} \rfloor$  and  $B = \lfloor (e^\alpha / 2)^{1/\beta} \rfloor + 1$ . Note that  $\alpha/\beta > \ln 2 (2/\beta + 1)$  since  $\alpha > 3 \ln 2$  and  $0 < \beta < 1$ . Hence  $(\lfloor (e^\alpha / 2)^{1/\beta} \rfloor + 1) (\lfloor (e^\alpha / 2)^{1/\beta} \rfloor) > \lfloor e^{\alpha/\beta} \rfloor \geq 2k$ , that is,  $f_D(k) > 0$ .

2. **Case 2:**  $k > \lfloor e^{\alpha/\beta} / 2 \rfloor$

$$f_D(k+1) \geq f_D(k) + 2k - 2d_{k+1} \geq f_D(k) \geq \dots \geq f_D(\lfloor e^{\alpha/\beta} / 2 \rfloor) > 0. \quad \square$$

## References

- [1] A. Ferrante, G. Pandurangan, K. Park, On the hardness of optimization in power-law graphs, Theoretical Computer Science 393 (2008) 220–230.
- [2] M. Faloutsos, P. Faloutsos, C. Faloutsos, On power-law relationships of the internet topology, in: Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, SIGCOMM '99, ACM, New York, NY, USA, 1999, pp. 251–262.
- [3] R. Albert, H. Jeong, A.L. Barabási, The diameter of the world wide web, Nature 401 (1999) 130–131.
- [4] S. Redner, How popular is your paper? An empirical study of the citation distribution, The European Physical Journal B - Condensed Matter and Complex Systems 4 (1998) 131–134.
- [5] S. Bornholdt, H.G. Schuster (Eds.), Handbook of Graphs and Networks: From the Genome to the Internet, John Wiley & Sons, Inc., New York, NY, USA, 2003.
- [6] S. Eubank, V.S.A. Kumar, M.V. Marathe, A. Srinivasan, N. Wang, Structural and algorithmic aspects of massive social networks, in: SODA '04, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2004, pp. 718–727.
- [7] I. Dinur, S. Safra, On the hardness of approximating minimum vertex cover, Annals of Mathematics 162 (2005).
- [8] G. Pandurangan, 2006. <https://sites.google.com/site/gopalpandurangan/papers-by-date>.
- [9] C. Gkantsidis, M. Mihail, A. Saberi, Conductance and congestion in power law graphs, SIGMETRICS Performance Evaluation Review 31 (2003) 148–159.
- [10] W. Aiello, F. Chung, L. Lu, A random graph model for massive graphs, in: STOC '00, ACM, New York, NY, USA, 2000, pp. 171–180.
- [11] W. Aiello, F. Chung, L. Lu, A random graph model for power law graphs, Experimental Mathematics 10 (2000) 53–66.
- [12] S. Janson, T. Luczak, I. Norros, Large cliques in a power-law random graph, 2009.
- [13] D. Kempe, J. Kleinberg, E. Tardos, Maximizing the spread of influence through a social network, in: KDD, ACM Press, 2003, pp. 137–146.
- [14] D. Kempe, J. Kleinberg, E. Tardos, Influential nodes in a diffusion model for social networks, in: ICALP, Springer Verlag, 2005, pp. 1127–1138.
- [15] P. Austrin, S. Khot, M. Safra, Inapproximability of vertex cover and independent set in bounded degree graphs, in: CCC '09, pp. 74–80.
- [16] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.
- [17] G. Bianconi, A.L. Barabási, Bose–Einstein condensation in complex networks, 2000.
- [18] I. Norros, H. Reittu, On a conditionally poissonian graph process, Advances in Applied Probability (2006) 38–59.
- [19] M. Chlebík, J. Chlebíková, Approximation hardness of dominating set problems in bounded degree graphs, Information and Computation 206 (2008) 1264–1275.
- [20] P. Alimonti, V. Kann, Hardness of approximating problems on cubic graphs, in: CIAC '97, Springer-Verlag, London, UK, 1997, pp. 288–298.
- [21] J. Hastad, Clique is hard to approximate within  $n^{1-\epsilon}$ , in: FOCS '96, IEEE Computer Society, Washington, DC, USA, 1996, p. 627.
- [22] V. Kann, On the approximability of NP-complete optimization problems, Ph.D. thesis, Royal Institute of Technology Stockholm, 1992.
- [23] P. Erdős, T. Gallai, Graphs with prescribed degrees of vertices, Matematikai Lapok 11 (1960) 264–274.
- [24] J. Bondy, U. Murty, Graph Theory with Applications, MacMillan, London, 1976.